Über die linearen Transformationen des komplexen Hilbertschen Raumes.

Von Friedrich Riesz in Szeged.

Einleitung.

Die Theorie der linearen Transformationen des Hilbertschen Raumes von unendlich vielen Dimensionen ist seit kurzer Zeit in einem neuen Aufschwunge begriffen. Hauptsächlich durch Problemstellungen und heuristische Ansätze der Quantenmechanik angeregt, untersucht man nun in aller Eile, inwiefern, in welcher Form und unter welchen Bedingungen die Hauptresultate der Hilbertschen Theorie der beschränkten Formen und der entsprechenden Transformationen auch im Unbeschränkten ihre Gültigkeit bewahren. Eines dieser Hauptresultate, vielleicht das wichtigste, nämlich der Satz über die Spektralzerlegung und die entsprechende Integraldarstellung der beschränkten quadratischen Formen, hat durch J. v. Neumann[1]) in einer auch weitere Zwecke verfolgenden Arbeit unter Verwendung von Ratschlägen von E. Schmidt eine unerwartet einfach lautende und in gewisser Hinsicht abschliessende Erweiterung erfahren, indem nämlich die Klasse jener Transformationen, für welche der Satz in entsprechender Form gültig bleibt, durch äusserst einfache Merkmale genau festgelegt wurde. Als wesentlichstes Hilfsmittel diente hiebei ein auf Cayley zurückgehender Gedanke, der es möglich macht, die Untersuchung der fraglichen Transformationen auf jene der sogenannten unitär orthogonalen (also beschränkten) Transformationen zurückzuführen.

---

[1]) J. v. Neumann, Allgemeine Eigenwerttheorie Hermitescher Funktionaloperatoren, *Math. Annalen,* 102 (1929), S. 49—131, insb. Kapitel IX und Anhang II. Der Anteil von E. Schmidt an den Untersuchungen ist daselbst in Fussnote [23]), S. 62, angegeben.

Für diese Transformationen spezieller Art gibt es, wie v. Neumann es zeigt, eine der Hilbertschen analoge Spektralzerlegung und Integraldarstellung, wobei das Spektrum statt auf der reellen Axe auf dem Einheitskreise gelegen ist; daraus ergibt sich mittels rechnerischer Umformung die Spektralzerlegung und Integraldarstellung der ursprünglichen Transformationen.

Eine andere, sehr übersichtliche Begründung derselben Tatsachen wurde durch M. H. Stone[2]) angedeutet; sie ist einer teilweise funktionentheoretischen Methode von T. Carleman zur Behandlung singulärer Integralgleichungen nachgebildet und auch mit der ursprünglichen Hilbertschen Abschnittsmethode verwandt.

Zweck der vorliegenden Arbeit ist, zu derselben Fragestellung in methodischer Hinsicht einiges beizutragen. Für den Hilbertschen Satz und die damit verbundenen Tatsachen haben seinerzeit E. Hellinger[3]) und der Verfasser[4]) je eine Methode ausgearbeitet, die beide unmittelbarer als der ursprüngliche Beweis ans Werk gehen. Ich musste mir daher nach Einsicht in die angeführten Arbeiten notwendigerweise die Frage aufwerfen, ob sich diese Methoden auch im Unbeschränkten bewähren. Meine Methode beruhte auf einer dem Matrizenkalkül ähnlichen Symbolik. Ich ordnete den Polynomen $P(t)=c_0+c_1t+\ldots+c_rt^r$ eines reellen Parameters t die symbolischen Polynome $P(A)=c_0E+c_1A+\ldots+c_rA^r$ der Transformation A zu, wo E die Identität, $A^2,\ldots, A^r$ die entsprechenden Iterierten von A bedeuteten und zeigte, dass das hiemit definierte Entsprechen zwischen $P(t)$ und $P(A)$ sich durch Grenzübergang auf eine allgemeine Klasse von Funktionen F übertragen lässt, woraus dann der Hilbertsche Satz durch Spezialisierung hervorging. Diese Methode umfasst die zeitlich vorangehende Methode von Hellinger, insofern nämlich letztere ebenfalls als eine derartige Zuordnung von $F(A)$ zu $F(t)$ gedeutet werden kann, wobei jedoch schon von

---

[2]) M. H. Stone, Linear transformations in Hilbert space, *Proceedings of the National Academy of Sciences, U. S. America,* **15** (1929), S. 198—200 und 423—425.

[3]) E. Hellinger, Neue Begründung der Theorie quadratischer Formen von unendlich vielen Veränderlichen, *Journal für die reine und angewandte Math.*, **136** (1909), S. 210—271.

[4]) F. Riesz, Über quadratische Formen von unendlich vielen Veränderlichen, *Nachrichten der k. Gesellschaft der Wissenschaften*, Göttingen, 1910, S. 190—195; *Les systèmes d'équations linéaires à une infinité d'inconnues* (Collection Borel), Paris 1913, Chapitre V.

Anfang an nur gewisse spezielle Funktionen betrachtet werden. Was nun meine Methode anbelangt, so scheint beim ersten Anblick ihre Anwendung auf nicht beschränkte Transformationen daran zu scheitern, dass in diesem Falle das Vorhandensein der Iterierten keineswegs evident ist. Für die Hellingersche Methode, die mit dem Umkehrproblem der Transformation $A - \mu E$ für nicht-reelle μ, also im wesentlichen mit der Zuordnung einer Transformation $F(A)$ zur Funktion $F(t) = (t-\mu)^{-1}$ ansetzt, fällt diese Schwierigkeit fort, indem nämlich für die durch v. Neumann betrachteten *hypermaximalen,* durch Stone als *„self-adjoint“* bezeichneten Transformationen die Existenz und die Beschränktheit der fraglichen Reziproken aus der Definition unmittelbar folgt und auch bei diesen Autoren als Ausgangspunkt für die weitere Untersuchung dient. Es liegt nun der Gedanke nahe, auch bei Anwendung meiner Methode *die Polynome durch lineare Verbindungen solcher Elementarbrüche und ihrer Potenzen, d. i. durch längs der unendlichen t-Axe beschränkte rationale Funktionen zu ersetzen,* wodurch dann die Anwendbarkeit der Methode tatsächlich gesichert ist und die in meinem Buche enthaltenen und teilweise im dritten Abschnitte der vorliegenden Arbeit wiederholten Ausführungen fast Wort für Wort auch unter den allgemeineren Voraussetzungen gelten.

Ich habe über diese Fragen vor einigen Monaten in Seminarvorträgen berichtet und ging daran, den Inhalt dieser Vorträge für den Druck niederzulegen. Dass nun die vorliegende Arbeit doch nicht genau diesem Plane entspricht, das hat seinen guten Grund. Während der Schriftlegung bemerkte ich nämlich, dass der eigentliche Kern sowohl der Hilbertschen Sätze, wie auch der Verallgemeinerungen, in einem sehr einfach lautenden, sich anstatt der kontinuierlichen Zerlegung nur auf eine sozusagen lokale Zerlegung beziehenden Satze steckt, der z. B. für eine beschränkte quadratische Form u. a. besagt, dass die Form, wenn sie nicht schon ursprünglich definit ist, *immer als Differenz zweier positiv definiten und zueinander orthogonalen Formen dargestellt werden kann.* Entsprechendes gilt für Hermite-Formen resp. Hermite-Transformationen; die gewünschte Zerlegung erhält man aus der schon angedeuteten Zuordnung von $F(A)$ zu $F(t)$, indem mar für $F_1(t)$ und $F_2(t)$ den positiven resp. negativen Teil der Funktion $F(t) = t$ wählt. Die Formulierung der Aussage im dritten Abschnitte weicht von der hier gegebenen insofern ab, dass zunächst die Identität E in zwei orthogonale Teile E_- und E_+ zerlegt und dann die gewünschte Zerlegung durch die Produkte AE_- und AE_+ geleistet wird, wodurch dann der Übergang zum Unbeschränkten nach Einführung der Reziproken von $A-iE$ sich unmittelbar durchführen lässt.

Ich habe mich deshalb entschlossen, meinen Weg, vielleicht sogar einen Umweg, über diesen lokalen Zerlegungssatz hindurch zu nehmen, weil von diesem Satze aus meines Erachtens nach der klarste Einblick in das enge Verhältnis zwischen den alten und den neuen Resultaten sich darbietet.

In den ersten beiden Abschnitten fasse ich das Notwendige über den komplexen Hilbertschen Raum und dessen beschränkte lineare Transformationen zusammen; im dritten komme ich bis zum soeben geschilderten Zerlegungssatz für beschränkte Transformationen; im vierten fasse ich die notwendigen Tatsachen über das Unbeschränkte zusammen und übertrage den Zerlegungssatz; die Spektralzerlegung und die entsprechenden Integraldarstellungen folgen im letzten Abschnitt.

1. Der komplexe Hilbertsche Raum.

Den folgenden Betrachtungen legen wir in der heute üblichen Weise den *abstrakten komplexen* Hilbert*schen Raum* $\mathfrak{H}$ zugrunde, von dem der Koordinatenraum von abzählbar unendlich vielen Dimensionen und der aus den in einer Menge von positivem Masse quadratisch integrierbaren Funktionen bestehende Funktionenraum je eine Verwirklichung sind.

Der Raum $\mathfrak{H}$ ist eine Mannigfaltigkeit von Elementen mit den folgende Eigenschaften.

Erstens ist $\mathfrak{H}$ *linear,* d. h. es gibt eine Addition $f+g$ der Elemente f, g und eine skalare Multiplikation cf mit komplexen Zahlen c, wobei für $c=0$ das Element cf unabhängig von f ist und mit 0 bezeichnet wird. Für diese Operationen gelten die Rechenregeln der Vektoralgebra.

Zweitens ist jedem Paare f, g eine bestimmte reelle oder komplexe Zahl, das *innere Produkt* (f, g), zugeordnet; für dieses gelten die Rechenregeln der komplexen Vektoralgebra, d. h. $(cf,g)=c(f,g)$; $(f_1+f_2, g)=(f_1,g)+(f_2,g)$; $(g,f)=\overline{(f,g)}$ und schliesslich $(f,f)\geqq 0$ und dann und nur dann 0, wenn $f=0$.

Durch $|f| = (f, f)^{1/2}$ ist der *Betrag* von f, durch $|f - g|$ die *Distanz* von f und g und damit eine *Metrik* in $\mathfrak{H}$ definiert.

Drittens ist $\mathfrak{H}$ *vollständig,* d. h. für eine Folge von Elementen $f_1, f_2, \ldots$ folgt aus $|f_m - f_n| \to 0$ für $m, n \to \infty$ die Existenz eines Elements f derart, dass $|f - f_n| \to 0$ für $n \to \infty$. Wir sagen kurz, f_n *konvergiert gegen* f und schreiben auch $f_n \to f$. Dass die genannte Bedingung auch notwendig ist und dass f durch die Folge $f_1, f_2, \ldots$ eindeutig bestimmt ist, schliesst man nach bewährtem Muster aus der Dreiecksungleichung $|f + g| \leqq |f| + |g|$, die samt der „Schwarzschen" Ungleichung $|(f, g)| \leqq |f|\,|g|$ auf bekannte Weise aus den postulierten Eigenschaften des inneren Produktes folgt. Mit Hilfe der letzten Ungleichung schliesst man noch, dass die Voraussetzung $f_n \to f$ die Limesgleichungen $(f_n, g) \to (f, g)$, $(g, f_n) \to (g, f)$ und die Voraussetzung $f_n \to f$, $g_n \to g$ die Limesgleichung $(f_n, g_n) \to (f, g)$ nach sich ziehen.

Viertens ist $\mathfrak{H}$ *separabel,* d. h. er enthält eine abzählbare, überall dichte Teilmenge. Der Ausdruck „überall dicht" besagt, wie gewöhnlich, dass jene Teilmenge für jedes Element f von $\mathfrak{H}$ Folgen $f_1, f_2, \ldots$ mit $f_n \to f$ enthält.

Fünftens enthält $\mathfrak{H}$ beliebig viele linear unabhängige Elemente, d. h. Elemente $f_1, \ldots, f_n$, für welche die Beziehung $c_1 f_1 + \ldots + c_n f_n = 0$ ($c_1, \ldots, c_n$ komplexe Zahlen) nur dann besteht, wenn $c_1 = \ldots = c_n = 0$ ist. Dieses letzte Postulat, das eine Dimensionsabgrenzung *nach unten* bewirkt, ist für uns unwesentlich, da unsere Betrachtungen *a fortiori* für jede endliche Dimension gelten.

Wir benötigen noch die folgenden beiden Existenzsätze, die man z. B. aus den entsprechenden Sätzen über den Koordinatenraum durch *Orthogonalisierung*[5]) der im vierten Postulat genannten abzählbaren Teilmenge und Einführung von Koordinaten gewinnt. Der erste der beiden Sätze kann übrigens auch schon aus den

---

[5]) „Orthogonal" besagt: $(f, g) = 0$, also auch $(g, f) = 0$; Orthogonalisierung ist das bekannte Verfahren von E. Schmidt, durch welches man aus jeder Folge von linear unabhängigen Elementen $f_1, f_2, \ldots$ eine neue Folge $f^{(1)}, f^{(2)}, \ldots$ bildet, deren Elemente paarweise orthogonal sind, den Betrag 1 haben und die Eigenschaft besitzen, dass man jedes f_n in der Form

$$f_n = c_1 f^{(1)} + \ldots + c_n f^{(n)}$$

darstellen kann, wo

$$c_k = (f_n, f^{(k)})$$

ist.

drei ersten Postulaten gefolgert werden, ist also auch von der im vierten Postulat geforderten Dimensionsabgrenzung *nach oben* unabhängig.

Satz **A.** *Ist eine lineare Mannigfaltigkeit* $\mathfrak{L}$ *aus* $\mathfrak{H}$, *d. i. eine Teilmenge, die mit* f *auch* cf, *mit* f *und* g *auch* $f+g$ *enthält, in* $\mathfrak{H}$ *nicht überall dicht, so gibt es ein Element* g *aus* $\mathfrak{H}$ *mit* $|g|=1$, *das zu allen Elementen aus* $\mathfrak{L}$ *orthogonal ist.*

Der zweite Satz bezieht sich auf Funktionen $l(f)$, die für sämtliche Elemente f von $\mathfrak{H}$ oder wenigstens für eine in $\mathfrak{H}$ überall dichte lineare Mannigfaltigkeit definiert sind. Eine solche Funktion $l(f)$ heisst *linear,* wenn sie *distributiv* und *beschränkt* ist, d. h. wenn $l(cf)=cl(f)$, $l(f+g)=l(f)+l(g)$ ist und die Gesamtheit der Werte $l(f)$ unter der Bedingung $|f|\leqq 1$ ebenfalls beschränkt ist.

Satz **B.** *Für jede lineare Funktion* $l(f)$ *gibt es ein eindeutig bestimmtes „erzeugendes" Element* g, *so dass*

$$l(f)=(f,g).$$ [6]

Umgekehrt erzeugt jedes Element g aus $\mathfrak{H}$ eine überall in $\mathfrak{H}$ definierte lineare Funktion $l(f)=(f,g)$.

2. Beschränkte lineare Transformationen.

Eine Transformation A, die jedem Elemente f aus $\mathfrak{H}$ ein bestimmtes Element Af desselben Raumes zuordnet, heisst *beschränkt linear,* wenn erstens $|Af|$ für alle f mit $|f|\leqq 1$ unterhalb einer Schranke liegt und zweitens A distributiv ist, d. h. $A(cf)=cAf$ und $A(f_1+f_2)=Af_1+Af_2$ gilt. Die genaue obere Schranke von $|Af|$ für $|f|\leqq 1$ heisst die *Schranke* von A und wird mit M_A bezeichnet.

Aus dieser Definition folgt unmittelbar, dass A *stetig* ist, d. h. dass mit $f_n \to f$ zugleich auch $Af_n \to Af$ besteht. Ferner folgt, dass für gegebenes g das innere Produkt (Af, g) eine lineare Funktion von f ist; nach Satz **B.** gibt es also ein eindeutig be-

---

[6]) Der Satz gilt allgemeiner für lineare Funktionen $l(f)$, die auf irgendeiner, also nicht notwendig überall dichten linearen Mannigfaltigkeit $\mathfrak{L}$ definiert sind; nur fällt dann das eindeutige Bestimmtsein von g hinweg. Die Eindeutigkeit kann in diesem Falle durch eine Zusatzforderung, wie z. B. dass $|g|$ möglichst klein ausfalle, oder auch, dass g der Mannigfaltigkeit $\mathfrak{L}$ resp. ihrer Derivierten angehöre, gesichert werden.

stimmtes g^* derart, dass

(1) $$(Af, g) = (f, g^*)$$

ist. Die durch (1) festgelegte Beziehung zwischen g und g^* ist wieder eine Transformation $g^* = A^* g$ und man zeigt nach wohlbekanntem Schema, dass auch A^* beschränkt linear und $M_{A^*} = M_A$ ist. A^* heisst die zu A *adjungierte* Transformation. Die Formel

$$(Af, g) = (f, A^* g),$$

die die Beziehung zwischen A und A^* festlegt, lässt sich auch in der Form

$$(A^* g, f) = (g, Af)$$

hinschreiben; daraus folgt, dass auch A zu A^* adjungiert ist.

Die Ungleichung

$$|(Af, g)| \leqq |Af|\,|g|$$

zeigt, dass auch der Betrag der sogenannten *Bilinearform* (Af, g) für $|f| \leqq 1$, $|g| \leqq 1$ unterhalb der Schranke M_A liegt. Setzt man speziell $g = Af/|Af|$, so ergibt sich, dass M_A *auch die genaue obere Schranke von* $|(Af, g)|$ *unter der genannten Bedingung ist.*

Ist identisch $(Af, g) = (f, Ag)$, also $A^* f = Af$ für alle f, kurz $A^* = A$, so heisst A *selbstadjungiert.* Die zugehörige Hermite-Form

$$(Af, f) = (f, Af)$$

ist dann reell und für $|f| \leqq 1$ ebenfalls beschränkt; die genaue untere und obere Schranke von (Af, f) für $|f| = 1$, m und M, heissen die *untere und obere Schranke der Form* (Af, f). Die Zahl M_A ist hier *gleich der grösseren unter den beiden Zahlen* $|m|$ *und* $|M|$, was man aus der Ungleichung

$$2|(Af, g) + (Ag, f)| = |(A(f+g), f+g) - (A(f-g), f-g)| \leqq \\ \leqq 2 \max(|m|, |M|) \{|f|^2 + |g|^2\}$$

erschliesst. Ersetzt man hier nämlich g durch ωg, wo

$$\omega = \operatorname{sgn}(Af, g) = \frac{(Af, g)}{|(Af, g)|}$$

und also $|\omega| = 1$ ist, so bleibt der rechtsstehende Ausdruck unverändert, während im linksstehenden beide Summanden gleich $|(Af, g)|$ werden, so dass also die Ungleichung

$$2|(Af, g)| \leqq \max(|m|, |M|) \{|f|^2 + |g|^2\}$$

und somit für $|f| \leqq 1$, $|g| \leqq 1$ die Ungleichung

$$|(Af, g)| \leqq \max(|m|, |M|)$$

folgt, woraus man die Ungleichung

$$M_A \leq \max(|m|, |M|)$$

abliest. Da andererseits für $|f| \leq 1$

$$|(Af, f)| \leq M_A$$

ist, so ist auch die obere Schranke von $|(Af, f)|$ für $|f| = 1$, d. i. $\max(|m|, |M|)$, höchstens gleich M_A. Also ist genau

$$M_A = \max(|m|, |M|), \tag{2}$$

w. z. b. w.

Ein wichtiges Korollar dieser Formel ist das folgende: *Ist $(Af, f) = 0$ für alle f, so ist $M_A = 0$ und also auch $Af = 0$ für alle f.*

Ersetzt man hier ferner A durch die Differenz $A - B$, wo A und B selbstadjungiert sind und $(Af, f) = (Bf, f)$ für alle f ist, so erhält man den Satz: *Ist für die beiden selbstadjungierten Transformationen A und B identisch $(Af, f) = (Bf, f)$, so ist auch identisch $Af = Bf$, d. i. $A = B$.*

Das *Produkt* AB zweier Transformationen A, B erklären wir, wie gewöhnlich, durch die Formel $ABf = A(Bf)$. Es ist offenbar $M_{AB} \leq M_A M_B$. Ferner folgt aus

$$(ABf, g) = (Bf, A^* g) = (f, B^* A^* g),$$

dass $(AB)^* = B^* A^*$ ist. A und B heissen *vertauschbar*, wenn $AB = BA$ ist. Sind A und B selbstadjungiert, so ist $(AB)^* = B^* A^* = BA$. *Das Produkt der selbstadjungierten Transformationen A, B ist also dann und nur dann selbstadjungiert, wenn $BA = AB$, d. i. wenn die Transformationen A, B vertauschbar sind.*

Wir sagen, *die Folge der Transformationen $A_1, A_2, \ldots$ konvergiere gegen die Transformation* A und schreiben $A_n \to A$, wenn erstens die Folge der Werte M_{A_n} beschränkt ist und zweitens $A_n f \to Af$ für alle f gilt.[7]) Offenbar folgt aus $A_n \to A$, dass auch $cA_n \to cA$, $BA_n \to BA$, $A_n B \to AB$. Allgemeiner folgt aus $A_n \to A$, $B_n \to B$, dass auch $A_n B_n \to AB$. Denn erstens ist die Folge der Werte $M_{A_n B_n} \leq M_{A_n} M_{B_n}$ beschränkt. Zweitens hat man

$$A_n B_n f = A_n B f - A_n (B - B_n) f,$$
$$A_n B f = A_n (Bf) \to A(Bf) = ABf;$$

---

[7]) Diese Art der Konvergenz einer Transformationsfolge nannte ich in meinem Buche starke Konvergenz. Da ich hier einzig und allein mit diesem Konvergenzbegriff arbeite, ist eine besondere Benennung überflüssig.

wir haben also nur zu zeigen, dass

$$A_n(B-B_n)f \to 0.$$

Nun hat man $B_n f \to Bf$, also $|Bf - B_n f| \to 0$ und daher auch

$$|A_n(B-B_n)f| \leqq M_{A_n} |Bf - B_n f| \to 0,$$

w. z. b. w.

Wenn $A_n \to A$ und die A_n selbstadjungiert sind, so folgt aus

$$(A_n f, g) = (f, A_n g), \ (A_n f, g) \to (Af, g), \ (f, A_n g) \to (f, Ag),$$

dass auch $(Af, g) = (f, Ag)$, dass also *auch A selbstadjungiert ist.*

Die selbstadjungierte Transformation A heisst *positiv definit,* wenn die Form (Af, f) es ist, d. i. wenn $(Af, f) \geqq 0$ ist für alle f, wenn also die zugehörige untere Schranke $m \geqq 0$ ist. A heisst *negativ definit,* wenn $-A$ positiv definit ist, wenn also $M \leqq 0$ ist. Wenn $A_n \to A$ und die A_n positiv (negativ) definit sind, so folgt aus

$$(A_n f, f) \to (Af, f),$$

dass *auch die Limestransformation A positiv (negativ) definit ist.*

3. Ein Zerlegungssatz.

In diesem Abschnitte soll ein Satz über die Zerlegung der selbstadjungierten beschränkten Transformation in definite Teile entwickelt werden, der im wesentlichen im Hilbertschen Satze über die Spektralzerlegung der beschränkten quadratischen Formen enthalten ist und den ich, wie schon gesagt, als den eigentlichen Kern dieses Satzes und der anschliessenden Tatsachen ansehe. Die folgenden Ausführungen sind im grossen und ganzen eine dem Zwecke angepasste Wiedergabe der entsprechenden Überlegungen aus meinem Buche. Mit E bezeichnen wir die Identität, mit A^n die iterierten Transformationen von A: $A^n = AA^{n-1}$, $A^1 = A$, $A^0 = E$.

Sei A eine selbstadjungierte beschränkte Transformation. Jedem Polynom

$$P(t) = c_0 + c_1 t + \ldots + c_r t^r$$

mit reellen Koeffizienten ordnen wir die Transformation

$$P(A) = c_0 E + c_1 A + \ldots + c_r A^r$$

zu; diese ist ebenfalls beschränkt und selbstadjungiert. Dieses Entsprechen zwischen $P(t)$ und $P(A)$ ist offenbar distributiv und multiplikativ, d. i. den aus den Polynomen $P_1(t)$, $P_2(t)$ gebildeten

Polynomen $a_1P_1(t)+a_2P_2(t)$, $P_1(t)\,P_2(t)$ entsprechen die Transformationen $a_1P_1(A)+a_2P_2(A)$ und $P_1(A)\,P_2(A)=P_2(A)\,P_1(A)$. Eine etwas tiefer liegende und für uns wichtige Eigenschaft des Entsprechens zwischen $P(t)$ und $P(A)$ besteht darin, dass *das Entsprechen von positivem Typus ist.* Dies soll folgendes bedeuten. Es seien m und M die untere und die obere Schranke der Form (Af,f); dann entspricht jedem im Intervall (m,M) positiven Polynom $P(t)$ eine positiv definite Transformation $P(A)$. Abweichend von meinem Buche, wo ich die entsprechende Tatsache aus der Hauptachsentransformation, also aus der Spektralzerlegung der gewöhnlichen quadratischen Formen mittels Grenzübergang herleite, will ich mich hier auf einen Satz von Hausdorff aus der elementaren Algebra stützen.

Wir können ohne Einschränkung der Allgemeinheit annehmen, dass $P(t)>0$ ist im abgeschlossenen Intervalle (m,M); denn ist nur $P(t)\geqq 0$, so können wir anstatt $P(t)$ das Polynom $P(t)+\varepsilon$ mit $\varepsilon>0$ betrachten und erhalten dann, dass

$$(P(A)f,f)+\varepsilon(f,f)\geqq 0$$

ist für jedes positive ε, dass also auch

$$(P(A)f,f)\geqq 0$$

und somit $P(A)$ positiv definit ist.

Es sei also $P(t)>0$ für $m\leqq t\leqq M$; es sei ferner $P(t)$ vom Grade r und es sei s eine ganze Zahl $\geqq r$. Ersetzt man dann jedes Glied c_kt^k durch den gleichwertigen Ausdruck

$$\frac{c_k}{(M-m)^s}\,[M(t-m)+m(M-t)]^k\,[(t-m)+(M-t)]^{s-k},$$

so erhält $P(t)$ nach Ausführung der Binomentwicklungen und nach Umordnung die Form

$$P(t)=\sum_{j=0}^{s} a_j\,(t-m)^j\,(M-t)^{s-j}. \tag{1}$$

Der Hausdorffsche Satz besagt nun, dass für genügend grosse Werte von s die Koeffizienten a_j dieser Entwicklung sämtlich positiv sind.[8]) Dann setzt also die Darstellung (1) die Positivität von $P(t)$ sozusagen in Evidenz. Wählen wir nun für s eine ungerade Zahl, so setzt auch die entsprechende Entwicklung von $P(A)$,

[8]) F. Hausdorff, Summationsmethoden und Momentfolgen I, *Math. Zeitschrift*, 9 (1921), S. 74—109, insb. S. 98—99.

$$P(A) = \sum_{j=0}^{s} a_j (A - mE)^j (ME - A)^{s-j}$$

den positiv definiten Charakter von $P(A)$ in Evidenz. Dann ist nämlich jedes einzelne Glied der rechtsstehenden Summe positiv definit. Denn die beiden Exponenten j, $s-j$ sind von verschiedener Parität; ist z. B. j eine gerade Zahl, so setze man

$$A_1 = (A - mE)^{\frac{j}{2}} (ME - A)^{\frac{s-j-1}{2}}; \quad A_2 = ME - A;$$

das entsprechende Glied ist dann, abgesehen vom positiven Zahlenfaktor, von der Form $A_1^2 A_2$; A_1 und A_2 sind vertauschbar und A_2 ist positiv definit. Also ist für jedes f

$$(A_1^2 A_2 f, f) = (A_2 g, g) \geqq 0 \qquad (g = A_1 f).$$

Ebenso erledigt man den entgegengesetzten Fall. Damit ist unsere Behauptung bewiesen.[9])

Wir sind nunmehr in der Lage, das Entsprechen zwischen Polynom und Transformation auf Funktionen allgemeinerer Art, darunter auch unstetige, zu erweitern.

Es sei $P_1(t), P_2(t), \ldots$ eine für $m \leqq t \leqq M$ monoton wachsende und beschränkte Polynomfolge, die also gegen eine beschränkte Funktion $F(t)$ konvergiert. Wir beweisen, dass dann auch *die Folge der Transformationen $P_r(A)$ konvergent ist.*

Wir dürfen annehmen, unsere Polynome seien für $m \leqq t \leqq M$ nichtnegativ; im entgegengesetzten Falle können wir dies z. B.

---

[9]) Die Richtigkeit der Behauptung ergibt sich auch in evidenter Weise, nämlich durch Zerlegung von $P(t)$ in lineare und quadratische Faktoren, aus dem folgenden Satz: *Sind die beiden selbstadjungierten Transformationen A und B positif definit und miteinander vertauschbar, so ist auch die Transformation $AB = BA$ positiv definit.* Ein elementarer Beweis dieses Satzes lässt sich wie folgt führen. Wir können annehmen, dass $M_A \leqq 1$; ferner setzen wir $A_1 = A$ und allgemein $A_{n+1} = A_n - A_n^2$. Dann sind sämtliche A_n und $E - A_n$ positiv definit. Dies folgt von n auf $n+1$ auf Grund der Formeln

$$A_{n+1} = A_n^2 (E - A_n) + A_n (E - A_n)^2, \quad E - A_{n+1} = (E - A_n) + A_n^2.$$

Da nun $A = A_1^2 + \ldots + A_n^2 + A_{n+1}$ ist, so hat man also

$$(A_1^2 f, f) + \ldots + (A_n^2 f, f) \leqq (Af, f);$$

daraus folgt aber die Konvergenz der Reihe $\Sigma (A_n^2 f, f)$ und also auch die Limesgleichung $|A_n f|^2 = (A_n^2 f, f) \to 0$. Somit ist schliesslich in evidenter Schreibweise $A = \Sigma A_n^2$, also auch $AB = \Sigma A_n^2 B = \Sigma A_n B A_n$. Damit ist AB in (unendlich viele) positiv definite Summanden zerlegt und somit auch selbst positiv definit.

durch Substraktion von $P_1(t)$ aus sämtlichen Polynomen erreichen; die Differenzen $P_s(A)-P_r(A)$, die für die folgenden Abschätzungen einzig in Betracht kommen, werden ja hiedurch nicht berührt. Unter der gemachten Annahme ist dann auch die Folge $P_1^2(t)$, $P_2^2(t), \ldots$ monoton wachsend und beschränkt. Sei $P_r(t) \leqq K$, also $P_r^2(t) \leqq K^2$. Dann sind die Transformationen $P_{r+1}^2(A)-P_r^2(A)$ und $K^2E-P_r^2(A)$, die ja nicht-negativen Polynomen entsprechen, positiv definit und also ist für ein jedes f die Folge der Werte $(P_r^2(A)f, f)$ wachsend und beschränkt; die Folge konvergiert somit gegen einen endlichen Grenzwert. Ferner ist für $r<s$

$$P_r^2(t) \leqq P_r(t)\, P_s(t) \leqq P_s^2(t);$$

daraus folgt wie vorhin, dass auch

$$(2) \qquad (P_r^2(A)f, f) \leqq (P_r(A)P_s(A)f, f) \leqq (P_s^2(A)f, f).$$

Also konvergieren für $r, s \to \infty$ alle drei Grössen (2) gegen denselben Grenzwert; daraus schliesst man aber, dass

$$((P_s(A)-P_r(A))^2 f, f) = |P_s(A)f - P_r(A)f|^2 \to 0,$$

dass also die Folge der Elemente $P_r(A)f$ konvergiert. Da ausserdem $M_{P_r(A)} \leqq K$ ist für alle r, so *konvergiert also die Folge der Transformationen $P_r(A)$ gegen eine wohlbestimmte beschränkte und selbstadjungierte Transformation, die wir der Grenzfunktion $F(t)$ zuordnen und dementsprechend mit $F(A)$ bezeichnen.*

Wir haben noch zu zeigen, dass diese Definition eine *eindeutige* ist, also dass $F(A)$ nur von $F(t)$ und nicht von der speziellen Wahl der approximierenden Polynome abhängt; d. i. wir haben zu zeigen, dass aus der Annahme

$$Q_1(t) \leqq Q_2(t) \leqq \ldots \to F(t) \qquad (m \leqq t \leqq M),$$

wo die $Q_r(t)$ ebenfalls Polynome sind, die Gleichung

$$\lim P_r(A) = \lim Q_r(A)$$

folgt. Um dies zu beweisen, hat man nur zu bemerken, dass, wie man z. B. auf Grund des Dinischen Satzes über monotone Folgen stetiger Funktionen oder auch durch Anwendung des Borelschen Überdeckungssatzes erkennt, für jedes r und für genügend grosse s die Ungleichungen

$$P_r(t) - {}^1\!/r \leqq Q_s(t), \quad Q_r(t) - {}^1\!/r \leqq P_s(t)$$

und damit auch die Ungleichungen

$$(P_r(A)f, f) - {}^1\!/r\,(f, f) \leqq (Q_s(A)f, f), \quad (Q_r(A)f, f) - {}^1\!/r\,(f, f) \leqq (P_s(A)f, f)$$

gelten. Aus den letzteren folgt durch Grenzübergang, wenn wir vorübergehend $P_r(A) \to P$, $Q_r(A) \to Q$ setzen, dass identisch für alle f

$$(Pf, f) = (Qf, f)$$

und somit $P = Q$ gilt. Damit ist die Eindeutigkeit von $F(A)$ gezeigt. Die eine Hälfte des Beweises liefert zugleich das folgende, allgemeiner lautende Resultat:

Ist für die beiden auf der Strecke $m \leqq t \leqq M$ monoton wachsenden und beschränkten Polynomfolgen $P_n(t)$, $Q_n(t)$

$$P_n(t) \to F(t), \quad Q_n(t) \to G(t),$$

ferner $F(t) \leqq G(t)$, dann ist auch

$$(F(A)f, f) \leqq (G(A)f, f), \tag{3}$$

(kurz $F(A) \leqq G(A)$), *d. h. die Transformation $G(A) - F(A)$ ist positiv definit.*

Die Bedeutung dieses Satzes wird besonders klar, wenn man das Entsprechen zwischen $F(t)$ und $F(A)$ auch auf die Differenzen der bisher betrachteten Funktionen ausdehnt, indem man der Funktion $F_1(t) - F_2(t)$ die Transformation $F_1(A) - F_2(A)$ zuordnet. Was zunächst die Eindeutigkeit dieser Definition betrifft, so hat man zu zeigen, dass aus

$$F_1(t) - F_2(t) = F_3(t) - F_4(t) \qquad (m \leqq t \leqq M)$$

die Gleichung

$$F_1(A) - F_2(A) = F_3(A) - F_4(A)$$

folgt. Dies wird sofort klar, wenn man die beiden Gleichungen auf die Form $F_1 + F_4 = F_2 + F_3$ bringt. Dann besagt ja die Behauptung nichts anderes, als die Additivität des Entsprechens für die schon betrachteten Funktionen und diese folgt unmittelbar aus der entsprechenden Tatsache für Polynome.

Für die neu betrachtete Funktionenklasse bedeutet die Formel (3) nichts anderes, als dass *das Entsprechen auch für diese allgemeine Klasse von positivem Typus ist,* d. i. dass jeder für $m \leqq t \leqq M$ nicht-negativen Funktion eine positiv definite Transformation entspricht.

Die Distributivität des Entsprechens ist auch für die neue Funktionenklasse evident. Was den multiplikativen Charakter des Entsprechens betrifft, so sind ja auch diese Funktionen $F(t)$ Limesfunktionen gewisser beschränkter Polynomfolgen $P_n(t)$, derart, dass $P_n(A) \to F(A)$ gilt und der multiplikative Charakter ergibt sich aus jenem für Polynome auf Grund des Satzes, nach welchem aus $A_n \to A$ und $B_n \to B$ die Limesgleichung $A_n B_n \to AB$ folgt.

Aus dem multiplikativen Charakter folgt speziell auch die Tatsache, dass *sämtliche Transformationen $F(A)$ miteinander vertauschbar sind.* Da ferner sämtliche Transformationen $P(A) = c_0 E + c_1 A + \ldots + c_r A^r$ offenbar mit jeder mit A vertauschbaren Transformation B, auch mit den nicht selbstadjungierten, vertauschbar sind, so folgt durch Grenzübergang die entsprechende Tatsache für sämtliche $F(A)$.

Noch eine letzte Bemerkung allgemeiner Art, die wir dann ebenso, wie die bisher bewiesenen Tatsachen, nur auf einen ganz speziellen Fall anzuwenden haben werden. Es bestehe für ein Element $f \neq 0$ und für den Wert λ die Gleichung $Af = \lambda f$, d. h. es sei, mit einer geläufigen Ausdrucksweise, $1/\lambda$ ein *Eigenwert* von A und f ein zugehöriges *Eigenelement.* Dann ist offenbar $P(A)f = P(\lambda)f$ und daher nach Grenzübergang $F(A)f = F(\lambda)f$. *Also ist $1/F(\lambda)$ ein Eigenwert der Transformation $F(A)$ und f ist ein zugehöriges Eigenelement.*

Wie gesagt, werden wir unsere Resultate nur auf einige spezielle Funktionen anzuwenden haben; es kommt uns daher nicht darauf an, die allgemeine Funktionenklasse, für welche wir das Entsprechen definiert haben, näher zu beschreiben. Es genügt für uns reichlich, wenn wir feststellen, dass *sämtliche für $m \leqq t \leqq M$ stetige Funktionen, wie auch die Limesfunktionen der aus solchen gebildeten beschränkten und monotonen Folgen in unserer Klasse enthalten sind.* Sei in der Tat $F(t)$ eine stetige Funktion und es sei $P_n(t)$ ein Polynom, welches die Funktion $F(t) - 1/2^n$ mit der Genauigkeit $1/2^{n+2}$ annähert; dann geht $P_n(t)$ wachsend gegen $F(t)$. Ist ferner $F(t)$ zwar unstetig, aber beschränkt und Limesfunktion einer wachsenden Folge von stetigen Funktionen $F_n(t)$, so wähle man für $P_n(t)$ ein Annäherungspolynom von der Genauigkeit $1/2^{n+2}$ der stetigen Funktion $F_n(t) - 1/2^n$.

Nun kommen wir zum angekündigten Zerlegungssatze. Unsere Klasse enthält offenbar die Funktionen $1, t$, dann die Funktion $E_-(t)$, die für $t < 0$ gleich 1, für $t \geqq 0$ gleich 0 ist, wie auch die Funktion $E_+(t) = 1 - E_-(t)$, die für $t \geqq 0$ den Wert 1 hat, sonst verschwindet, schliesslich die aus diesen Funktionen gebildeten Produkte, also die Funktionen t^- und t^+, deren erste für $t < 0$, die zweite für $t > 0$ gleich t sind und sonst verschwinden.

Die Transformationen $E_-(A)$, $E_+(A)$ sind selbstadjungiert und ihren eigenen Iterierten gleich, da ja die entsprechenden Funktionen nur die Werte 0 und 1 annehmen und somit mit ihren Potenzen zusammenfallen. Selbstadjungierte Transformationen mit dieser Eigenschaft werden wir in Anschluss an eine Benennung bei HILBERT als *Einzeltransformationen* bezeichnen.[10]) Ist P eine solche, so ist wegen $P^2 = P$ auch $(Pf, f) = (P^2 f, f) = (Pf, Pf) = |Pf|^2 \geqq 0$ und *also ist P positiv definit.*

Die betrachteten Transformationen sind, wie alle $F(A)$, mit A selbst und mit jeder mit A vertauschbaren Transformation vertauschbar. Ferner entsprechen die Transformationen $AE_-(A)$, $AE_+(A)$ den Funktionen t^-, t^+, deren erste nirgends positiv, die zweite nirgends negativ ist, also ist die Transformation $AE_-(A) = E_-(A)A$ negativ, die Transformation $AE_+(A) = E_+(A)A$ positiv definit.[11]) Ausserdem ist $E_-(A) + E_+(A) = E$ und $AE_-(A) + AE_+(A) = A$. Endlich folgt aus $E_-(0) = 0$, $E_+(0) = 1$, dass alle f, für welche $Af = 0$ ist, die also zu dem Eigenwerte $\lambda = \infty$ ($^1/\lambda = 0$) gehörige Eigenelemente von A sind, auch Eigenelemente von $E_-(A)$ und $E_+(A)$ u. zw. in Bezug auf die Eigenwerte ∞ resp. 1 sind, dass also für diese Elemente $E_-(A)f = 0$, $E_+(A)f = f$ gilt.

Zusammenfassend haben wir den Satz:

Zu jeder selbstadjungierten beschränkten Transformation A gibt es eine Zerlegung der identischen Transformation E in zwei Einzeltransformationen: $E = E_- + E_+$ derart, dass

1° *diese Transformationen mit A und mit jeder mit A vertauschbaren beschränkten linearen Transformation vertauschbar sind;*

2° *die Transformationen AE_- und AE_+ negativ resp. positiv definit sind und ihre Summe gleich A ist;*

3° *dass für alle f, für welche $Af = 0$ ist, auch $E_- f = 0$, $E_+ f = f$ gelten.*

---

[10]) J. v. NEUMANN nennt sie *Projektionsoperatoren*. Diese Benennung ist dadurch begründet, dass Pf und $f - Pf$ zueinander orthogonal stehen und somit Pf als die orthogonale Projektion von f auf die aus $\mathfrak{H}$ durch P entstehende lineare Mannigfaltigkeit gedeutet werden kann.

[11]) Die beiden Transformationen $A_- = AE_-(A)$, $A_+ = AE_+(A)$ kann man auch unmittelbar, ohne Heranziehung unstetiger $F(t)$, als die den Funktionen t^-, t^+ zugeordnete Transformationen einführen; die Transformationen $E_-(A)$, $E_+(A)$ lassen sich dann nachträglich als jene Transformationen erklären, welche sämtliche $A_- f$ in sich resp. in 0 und die zu allen diesen orthogonalen Elemente in 0 resp. in sich überführen.

4. Die allgemeine selbstadjungierte Transformation.

Die Verallgemeinerung des Begriffes der selbstadjungierten Transformation durch v. Neumann und Schmidt setzt damit ein, dass man die Forderung der Beschränkheit fallen lässt und auch davon absieht, dass die Transformation A für $\mathfrak{H}$ ausnahmslos Sinn habe, wie dies ja z. B. für den Raum der quadratisch integrierbaren Funktionen und für die Differentiation als Transformation nicht mehr der Fall ist. Man begnügt sich damit, dass A *für eine in* $\mathfrak{H}$ *überall dichte lineare Mannigfaltigkeit* $\mathfrak{L}$ *definiert sei* (die natürlich auch mit $\mathfrak{H}$ zusammenfallen darf) und man fordert auch weiterhin die Distributivität. Nun kommt alles darauf an, den Begriff der adjungierten Transformation festzulegen. Das innere Produkt (Af, g) ist für alle f aus $\mathfrak{L}$ und alle g aus $\mathfrak{H}$ definiert; für ein gegebenes g ist $l(f) = (Af, g)$ eine distributive Funktion von f; ist diese für $|f| \leq 1$ auch noch beschränkt, dann gibt es nach Satz **B.** ein bestimmtes Element g^* aus $\mathfrak{H}$ derart, dass für alle f aus $\mathfrak{L}$

$$(Af, g) = (f, g^*)$$

ist. Auf diese Weise wird dem Element g ein eindeutig bestimmtes Element g^* zugeordnet. Die Elemente g, für welche dies der Fall ist, bilden offenbar eine lineare Mannigfaltigkeit $\mathfrak{L}^*$ und die Zuordnung ist offenbar distributiv. Ist nun $\mathfrak{L}^*$ *in* $\mathfrak{H}$ *überall dicht,* so ist also durch $A^*g = g^*$, d. i. durch die Formel

$$(Af, g) = (f, A^*g)$$

eine Transformation von der betrachteten Art definiert. Wir nennen A^* die zu A *adjungierte* Transformation. Gehört zu A eine adjungierte Transformation A^*, so gehört auch zu A^* eine Adjungierte A^{**}, die gewiss wenigstens für alle f aus $\mathfrak{L}$ definiert ist und für diese mit A zusammenfällt, im allgemeinen aber eine Fortsetzung von A definiert. Die Adjungierte von A^{**} ist dann wieder gleich A^*. Ist $A = A^*$ und also auch $A = A^{**}$, so sagen wir, A sei eine *selbstadjungierte* Transformation.[12]) Für eine solche ist $(Af, g) = (f, Ag)$

[12]) J. v. Neumann nennt diese Transformationen *hypermaximal;* die Benennung ist durch die Anlage seiner Arbeit geboten, in der er nämlich die diversen Fortsetzungmöglichkeiten der sog. Hermiteschen Operatoren untersucht. Hätte man den Namen Hermite nicht vorzeitig zur Benennung eines allzu allgemeinen Begriffs herangezogen, so wären nach meiner Ansicht die selbstadjungierten Transformationen als Hermitesche zu bezeichnen, da sie die naturgemässe Verallgemeinerung der gleichbenannten Matrizen bilden.

für alle f, g aus $\mathfrak{L}$ und speziell $(Af, f) = (f, Af)$ für alle f aus $\mathfrak{L}$, woraus die Realität der HERMITE-Form (Af, f) folgt.

Die selbstadjungierte Transformation A heisst auch hier positiv (negativ) definit, wenn (Af, f) ständig $\geqq 0$ ($\leqq 0$) ausfällt.

Die Einordnung der beschränkten linearen Transformationen unter diese allgemeine Begriffsbildungen liegt an der Hand. Ist eine Transformation A von der betrachteten Art, die selbstadjungiert oder nur eine Adjungierte ist, beschränkt, d. h. ist für $|f| \leqq 1$ auch $|Af|$ beschränkt, dann ist A notwendigerweise für ganz $\mathfrak{H}$ definiert und ist also eine beschränkte lineare Transformation im gewohnten Sinne. Bei selbstadjungierten Transformationen ist die Beschränktheit gleichwertig mit dem Vorhandensein von Zahlen m, M derart, dass die Transformationen $A - mE$, $ME - A$ positiv definit ausfallen.

Eine Transformation A, die eine Adjungierte A^* zulässt, ist sicher in ihrem ganzen Verlaufe beschränkt, d. i. ihre natürliche Fortsetzung A^{**} ist für ganz $\mathfrak{H}$ definiert und beschränkt, sobald A für irgendeine in $\mathfrak{H}$ überall dichte lineare Mannigfaltigkeit $\mathfrak{L}_0$ beschränkt ist, wenn es also eine Zahl M_0 gibt, so dass $|Af| \leqq M_0$ für alle f mit $|f| \leqq 1$ aus $\mathfrak{L}_0$. Denn aus dieser Annahme folgt $|(f, A^*g)| = |(Af, g)| \leqq M_0$ für alle f mit $|f| \leqq 1$ aus $\mathfrak{L}_0$ und alle g mit $|g| \leqq 1$, für welche A^*g Sinn hat. Zufolge der Stetigkeit des inneren Produktes bleibt die Ungleichung $|(f, A^*g)| \leqq M_0$ auch für alle f mit $|f| \leqq 1$ aus $\mathfrak{H}$, also speziell für $f = A^*g / |A^*g|$ bestehen; durch Einsetzen folgt, dass $|A^*g| \leqq M_0$ für alle g mit $|g| \leqq 1$, für welche nur A^*g Sinn hat. Somit ist A^* beschränkt und also sind auch die Funktionen $l^*(g) = (f, A^*g)$ beschränkt, nämlich $|(f, A^*g)| \leqq M_0 |f|\, |g|$, und somit linear für alle f aus $\mathfrak{H}$. Daraus folgt unmittelbar, dass *A^{**} für ganz $\mathfrak{H}$ vorhanden und beschränkt ist, u. zw. ist $M_{A^{**}} \leqq M_0$.*

Ein wichtiger Spezialfall ist der, wo $M_0 = 0$, d. i. $Af = 0$ für alle f aus $\mathfrak{L}_0$. Dann ist auch $M_{A^{**}} = 0$, also *identisch* $A^{**} = 0$.

Ist A *selbstadjungiert,* dann kann man schon auch aus der Beschränktheit von (Af, f) für f mit $|f| \leqq 1$ aus $\mathfrak{L}_0$ auf jene von A schliessen. Genauer gesagt: *sind m_0, M_0 die untere und obere Schranke von (Af, f) auf $\mathfrak{L}_0$ bei der Bedingung $|f| = 1$, so ist A beschränkt und es ist $m = m_0$, $M = M_0$.* Zum Beweise nehmen wir an, dass $m_0 = -M_0$ ist; der allgemeine Fall lässt sich ja offenbar auf diesen Fall zurückführen, indem man nämlich an

Stelle von A die Transformation $A - \frac{m_0 + M_0}{2} E$ betrachtet. Bei der speziellen Annahme ist $\max(|m_0|, |M_0|) = M$ und es folgt durch Anwendung desselben Gedankenganges, durch welchen wir seinerzeit die Beziehung zwischen m, M und M_A feststellten, jedoch nur auf $\mathfrak{L}_0$ statt auf $\mathfrak{H}$, dass auch

(1) $$|(Af, g)| \leqq M_0 |f| \, |g|$$

für f, g aus $\mathfrak{L}_0$. Da man aber Af durch g aus $\mathfrak{L}_0$ beliebig approximieren kann, so folgt aus (1) durch Grenzübergang

$$|(Af, Af)| \leqq M_0 |f| \, |Af|$$

und also $|Af| \leqq M_0 |f|$ für f aus $\mathfrak{L}_0$. Also ist die Transformation $A = A^{**}$ für ganz $\mathfrak{H}$ sinnvoll und beschränkt, u. zw. ist $M_A \leqq M_0$. Speziell ist $|(Af, f)| \leqq M_0 |f|^2$, d. i.

$$(Af, f) + M_0 (f, f) \geqq 0, \quad M_0 (f, f) - A(f, f) \geqq 0$$

und somit $m \geqq -M_0$, $M \leqq M_0$; da anderseits $m \leqq -M_0$, $M \geqq M_0$ gelten muss, so folgt $m = -M_0$, $M = M_0$, w. z. b. w.

Ist speziell $(Af, f) = 0$ für alle f aus $\mathfrak{L}_0$, so ist $m_0 = M_0 = 0$, also auch $m = M = 0$ und daher auch $M_A = 0$, d. i. $A = 0$.

Es sei nun A eine beschränkte oder nicht beschränkte, selbstadjungierte Transformation; f sei ein gegebenes Element aus $\mathfrak{H}$. Wir betrachten die Gleichung

(2) $$Ax - ix = f \qquad (i = \sqrt{-1})$$

mit dem unbekannten Element x und zeigen, dass (2) *eine und nur eine Lösung zulässt.*

Erstens, im homogenen Fall, d. i. für $f = 0$, ist $x = 0$ die einzige Lösung; denn ist x eine Lösung, so hat man

$$(Ax, x) - i(x, x) = (Ax - ix, x) = (0, x) = 0$$

und somit

$$(Ax, x) = i(x, x),$$

woraus wegen der Realität von (Ax, x) und von (x, x) folgt, dass $|x|^2 = (x, x) = 0$, also $x = 0$ sein muss.

Zweitens durchlaufe x die lineare Mannigfaltigkeit $\mathfrak{L}$, für welche A Sinn hat; wir haben zu zeigen, dass dann $f = Ax - ix$ ganz $\mathfrak{H}$ durchläuft. Jedenfalls bilden die entsprechenden f eine lineare Mannigfaltigkeit $\mathfrak{L}_1$ und diese ist in $\mathfrak{H}$ überall dicht; denn sonst gäbe es ein Element g mit $|g| = 1$, die zu sämmtlichen f orthogonal wäre; es wäre daher

$$(Ax, g) = (ix, g) = (x, -ig)$$

für alle x aus $\mathfrak{L}$ und wegen $A^* = A$ hätte man also $Ag = -ig$ und somit

$$(Ag, g) = (-ig, g) = -i(g, g) = -i,$$

was ja der Realität von (Ag, g) widerspricht.

Die Mannigfaltigkeit $\mathfrak{L}_1$ ist somit in $\mathfrak{H}$ überall dicht; um zu beweisen, dass die beiden identisch sind, haben wir noch zu zeigen, dass für jede konvergente Folge $f_n \to f$ mit Elementen aus $\mathfrak{L}_1$ auch das Grenzelement f in $\mathfrak{L}_1$ enthalten ist. Mit anderen Worten haben wir aus

$$Ax_n - ix_n = f_n; \quad f_n \to f$$

auf die Existenz eines x mit

$$Ax - ix = f$$

zu schliessen.

Das machen wir so. Aus

$$\begin{aligned}|f_m - f_n|^2 &= |A(x_m - x_n) - i(x_m - x_n)|^2 = \\ &= (A(x_m - x_n) - i(x_m - x_n), A(x_m - x_n) - i(x_m - x_n)) = \\ &= (A(x_m - x_n), A(x_m - x_n)) + (x_m - x_n, x_m - x_n) = \\ &= |A(x_m - x_n)|^2 + |x_m - x_n|^2 \geqq |x_m - x_n|^2,\end{aligned}$$

also

$$|x_m - x_n| \leqq |f_m - f_n|,$$

folgt zunächst die Existenz eines x mit $x_n \to x$. Ferner hat man für jedes g aus $\mathfrak{L}$

$$(g, Ax_n) = (Ag, x_n) \to (Ag, x), \quad (g, Ax_n) = (g, f_n + ix_n) \to (g, f + ix);$$

somit ist für alle g aus $\mathfrak{L}$

$$(Ag, x) = (g, f + ix).$$

Dies besagt aber, zufolge der Selbstadjungiertheit von A, dass Ax sinnvoll und gleich $f + ix$ ist. Mit anderen Worten: x befriedigt die Gleichung (2), w. z. b. w.

Dass die Gleichung (2) für jedes f nur eine einzige Lösung zulässt, folgt in bekannter Weise aus der entsprechenden Tatsache bezüglich der homogen Gleichung.

Ordnet man nun jedem f die entsprechende Lösung x zu, so ist hiedurch eine beschränkte und zufolge der Eindeutigkeit der Lösung auch distributive, also eine beschränkte lineare Transformation B definiert. Diese ist mit der Transformation $A - iE$

durch die Beziehung

$$B(A-iE)=(A-iE)B=E \tag{3}$$

verknüpft, die man folgendermassen zu deuten hat. Einerseits macht B die Transformation $A-iE$ rückgängig für jedes f, für welches A (also auch $A-iE$) Sinn hat; andererseits wird B durch $A-iE$ rückgängig gemacht u. zw. für jedes f aus $\mathfrak{H}$. Indem wir uns einer gewohnten Ausdrucksweise bedienen, haben wir also den Satz:

Für jede selbstadjungierte lineare Transformation A besitzt die Transformation $A-iE$ eine beschränkte Reziproke.

Für den Fall beschränkter Transformationen hat dieser Satz, der sich in diesem Falle unmittelbar aus einem Satze von Toeplitz[13]) ergibt, als Ausgangspunkt für die Hellingersche Methode gedient. Im allgemeinen Falle scheint er für die Gewinnung der Spektralzerlegung zur Zeit unentbehrlich zu sein; er hat den genannten Untersuchungen und soll nun auch uns als Grundlage dienen.

Zwischen A und B besteht u. a. die folgende Beziehung:

Damit A für f Sinn habe, ist notwendig und hinreichend, dass f von der Form Bg sei, wo g ein beliebiges Element von $\mathfrak{H}$ bezeichnet.

Mit anderen Worten: B überführt $\mathfrak{H}$ genau in die lineare Mannigfaltigkeit $\mathfrak{L}$, die den Definitionsbereich von A bildet.

Hat nämlich Af Sinn, dann folgt aus (3), dass

$$f=Bg,\ g=Af-if$$

ist. Umgekehrt folgt für jedes $f=Bg$, dass

$$(A-iE)f=g;$$

also hat $A-iE$ und damit auch A für $f=Bg$ Sinn.

Ein zweiter wichtiger Zusammenhang zwischen A und B betrifft die Frage der Vertauschbarkeit von A mit einer beschränkten linearen Transformation C. Wir sagen mit J. v. Neumann, die Transformationen A und C seien *vertauschbar,* wenn mit Af zugleich auch $ACf=A(Cf)$ Sinn hat und $ACf=CAf$ ist.[14]) Es besteht nun der Satz:

---

[13]) O. Toeplitz, Die Jacobische Transformation der quadratischen Formen von unendlich vielen Veränderlichen, *Nachrichten der k. Gesellschaft der Wissenschaften,* Göttingen, 1907, S. 101—109.

[14]) J. v. Neumann, Zur Algebra der Funktionaloperationen und der Theorie der normalen Operatoren, *Math. Annalen,* **102** (1929), S. 370—427, insb. S. 404.

Notwendig und hinreichend für die Vertauschbarkeit von C mit A ist die Vertauschbarkeit von C mit B.

Ist nämlich erstens C mit A vertauschbar, so ist sie es auch mit $A-iE$ und es ist infolge (3) und da A für alle Bg und somit auch für CBg Sinn hat,

$$BC = BC(A-iE)B = B(A-iE)CB = CB.$$

Ist zweitens C mit B vertauschbar, dann ist

$$C(A-iE)B = C = (A-iE)BC = (A-iE)CB$$

für alle g aus $\mathfrak{H}$, also ist $(A-iE)C$ für alle Bg, d. i. für alle f aus $\mathfrak{L}$ erklärt und es ist $C(A-iE)f = (A-iE)Cf$. Also ist C mit $A-iE$ und somit auch mit A vertauschbar.

Aus dem soeben bewiesenen Satze folgt unmittelbar, dass, *wenn $C_n \to C$, wo C und C_n beschränkte Transformationen bedeuten und wenn alle C_n mit A vertauschbar sind, dies auch für C der Fall ist.* Denn wegen $BC_n \to BC$, $C_nB \to CB$, $BC_n = C_nB$ ist dann auch $BC = CB$, also ist C mit B und daher auch mit A vertauschbar.[15])

Ist C mit A vertauschbar, so werden wir unter $CA = AC$ jene Transformation verstehen, die durch die Vorschrift $ACf = A(Cf)$ für alle jene f erklärt wird, für welche der letzte Ausdruck Sinn hat. Es kann z. B. AC für ganz $\mathfrak{H}$ erklärt sein und eine beschränkte Transformation darstellen, ohne dass dies für A selbst der Fall wäre, man denke nur an die Transformation $AB = E + iB$.

Ist speziell die mit A vertauschbare beschränkte Transformation C ebenfalls selbstadjungiert und ist ausserdem AC beschränkt, so ist auch AC selbstadjungiert. Denn dann besteht die Gleichung

$$(ACf, g) = (Cf, Ag) = (f, CAg) = (f, ACg)$$

gewiss für alle f aus $\mathfrak{H}$ und für alle g, für welche A Sinn hat. Also stimmt AC mit $(AC)^*$ für jene g überein, und da mit AC zugleich auch $(AC)^*$ beschränkt ist, so folgt daraus das Übereinstimmen der beiden Transformationen für ganz $\mathfrak{H}$.

Schliesslich führen wir noch die offenbar ebenfalls vorhandene Reziproke von $A + iE$ ein, die wir sofort mit B^* bezeichnen;

---

[15]) Dasselbe folgt auch aus der leicht beweisbaren Tatsache, dass die Vertauschbarkeit von C mit A gleichwertig ist mit dem Bestehen der Gleichung $(Cf, Ag) = (CAf, g)$ für alle f und g, für welche A Sinn hat.

denn es ist für alle f und g aus $\mathfrak{H}$

$$(Bf, g) = (Bf, (A+iE)B^*g) = (Bf, AB^*g) - i(Bf, B^*g) =$$
$$= (ABf, B^*g) - i(Bf, B^*g) = ((A-iE)Bf, B^*g) = (f, B^*g)$$

und somit ist B^* die Adjungierte von B, was wir schon mit der Schreibweise andeuteten. Da B^* ebenso wie B mit A vertauschbar ist, so ist sie es auch mit B.[16]) Wir betrachten nun die Transformation

$$D = B + B^*;$$

diese ist beschränkt, selbstadjungiert und mit A, B und B^* vertauschbar. Aus

$$B = B^*(A+iE)B; \quad B^* = B^*(A-iE)B$$

erhält man durch Addition eine zweite Darstellung

$$D = 2B^*AB.$$

Wenden wir auf D den Zerlegungssatz an; seien E_- und E_+ die entsprechenden Einzeltransformationen. Ich behaupte, dass diese auch für A eine Zerlegung derselben Art bestimmen. Erstens sind diese mit jeder mit A vertauschbaren beschränkten linearen Transformation vertauschbar. Denn ist C eine solche, so ist sie nach dem oben bewiesenen Satze auch mit B und ebenso mit B^*, also auch mit D und folglich mit E_- und E_+ vertauschbar. Speziell sind also E_- und E_+ mit B und daher auch mit A vertauschbar. Zweitens ist für alle f und g, für welche A Sinn hat und die somit auf die Form $f = Bf_1$, $g = Bg_1$ gebracht werden können,

$$(E_-Af, g) = (E_-ABf_1, Bg_1) = (B^*E_-ABf_1, g_1) = {}^1/_2(E_-Df_1, g_1) =$$
$$= {}^1/_2(f_1, E_-Dg_1) = (f_1, B^*E_-ABg_1) = (f, E_-Ag);$$

diese Formel zeigt, dass E_-A selbstadjungiert ist; setzt man speziell $g = f$, also $g_1 = f_1$, so zeigt die Formel noch, dass

$$(E_-Af, f) = {}^1/_2(E_-Df_1, f_1),$$

dass also mit E_-D auch E_-A negativ definit ist. Ebenso sieht man, dass E_+A selbstadjungiert und positiv definit ist. Endlich ist, wenn $Af = 0$, auch $Df = 2B^*BAf = 0$ und also $E_-f = 0$, $E_+f = f$.

Damit ist bewiesen, dass *der am Schlusse des letzten Abschnittes stehende Zerlegungssatz auch für die allgemeine selbstadjungierte Transformation A ohne sonstige Änderung gültig bleibt.*

---

[16]) Daraus folgt, dass für beschränktes C mit C zugleich auch C^* mit A vertauschbar ist. Denn aus $AC = CA$ folgt $BC = CB$, also auch $B^*C^* = C^*B^*$ und somit $AC^* = C^*A$, da ja in dem bezüglichen Satze offenbar B durch B^* ersetzt werden darf.

5. Die Spektralschar und die entsprechenden Integraldarstellungen.

Es sei A eine beliebige, beschränkte oder nicht beschränkte selbstadjungierte Transformation; dann sind bei reellem λ auch alle Transformationen $A-\lambda E$ setbstadjungiert Zu jeder dieser Transformationen gehören im Sinne des Zerlegungssatzes zwei Einzeltransformationen E_- und $E_+ = E - E_-$; wegen der Abhängigkeit dieser Transformationen vom Parameter λ werden wir sie von nun an mit E_λ resp. $E-E_\lambda$ bezeichnen.

Die auf diese Weise eingeführte Schar von Einzeltransformationen E_λ hat gewisse einfache Eigenschaften. *Erstens sind alle E_λ miteinander vertauschbar.* Denn E_{λ_1} ist vertauschbar mit $A-\lambda_1 E$, also auch mit $A-\lambda_2 E$; auf Grund des Zerlegungssatzes ist also E_{λ_2} mit E_{λ_1} vertauschbar. *Zweitens sei $\lambda_1 < \lambda_2$; dann ist $E_{\lambda_1} E_{\lambda_2} = E_{\lambda_1}$.* Diese Behauptung ist gleichwertig mit $E_{\lambda_1}(E-E_{\lambda_2}) = 0$. Um letztere zu beweisen, setzen wir $E_{\lambda_1}(E-E_{\lambda_2}) = (E-E_{\lambda_2}) E_{\lambda_1} = P$; dann ist

$$E_{\lambda_1} P = E_{\lambda_1}^2 (E-E_{\lambda_2}) = E_{\lambda_1}(E-E_{\lambda_2}) = P$$

und ebenso auch

$$(E-E_{\lambda_2}) P = P.$$

Es sei nun f ein Element, für welches A Sinn hat und es sei $Pf = g$. Nach dem Zerlegungssatze ist

$$(E_{\lambda_1}(A-\lambda_1 E) g, g) \leqq 0, \quad ((E-E_{\lambda_2})(A-\lambda_2 E) g, g) \geqq 0,$$

somit wegen

$$E_{\lambda_1} g = E_{\lambda_1} Pf = Pf = g \text{ und } (E-E_{\lambda_2}) g = (E-E_{\lambda_2}) Pf = Pf = g$$

auch

$$((A-\lambda_1 E) g, g) = ((A-\lambda_1 E) g, E_{\lambda_1} g) = (E_{\lambda_1}(A-\lambda_1 E) g, g) \leqq 0$$

und ebenso

$$((A-\lambda_2 E) g, g) \geqq 0,$$

woraus durch Substraktion

$$(\lambda_2-\lambda_1)(g, g) = (\lambda_2-\lambda_1) |g|^2 \leqq 0,$$

also $Pf = g = 0$ folgt. Es ist somit $Pf = 0$ für eine in $\mathfrak{H}$ überall dichte Mannigfaltigkeit; da P beschränkt ist, so folgt daraus $P = 0$, w. z. b. w.

Drittens ist für $\lambda_1 < \lambda_2$ die Differenz $E_{\lambda_2} - E_{\lambda_1}$ eine Einzeltransformation und somit auch positiv definit. Denn es ist nach dem soeben bewiesenen

$$(E_{\lambda_2}-E_{\lambda_1})^2=E_{\lambda_2}^2-2E_{\lambda_1}E_{\lambda_2}+E_{\lambda_1}^2=E_{\lambda_2}-2E_{\lambda_1}+E_{\lambda_1}=E_{\lambda_2}-E_{\lambda_1}.$$

Die Definitheit von $E_{\lambda_2}-E_{\lambda_1}$ besagt, anders ausgedrückt, dass *die Form*

$$(E_\lambda f, f)=(E_\lambda f, E_\lambda f)=|E_\lambda f|^2$$

für jedes f eine monoton wachsende Funktion von λ darstellt. Wegen

$$|f|^2-|E_\lambda f|^2=(f,f)-(E_\lambda f,f)=((E-E_\lambda)f,f)\geq 0$$

bleibt der Wertevorrat der obigen Form zwischen 0 und $|f|^2$.

Lassen wir jetzt λ gegen $-\infty$ resp. ∞ gehen; wir werden zeigen, dass in dem ersten Falle $E_\lambda \to 0$, im zweiten $E_\lambda \to E$ gilt. Es genügt, den ersten Fall zu betrachten; der zweite wird analog erledigt, indem man E_λ durch $E-E_\lambda$ ersetzt. Sei f ein Element, für welches Af Sinn hat; dann ist für $\lambda<0$

$$|E_\lambda f|^2=(E_\lambda f,f)={}^1/_\lambda(E_\lambda Af,f)-{}^1/_\lambda(E_\lambda(A-\lambda E)f,f)\leq$$
$$\leq {}^1/_\lambda(E_\lambda Af,f)\leq {}^1/_{|\lambda|}\,|E_\lambda Af|\,|f|\leq {}^1/_{|\lambda|}\,|Af|\,|f|$$

und also $E_\lambda f\to 0$ für $\lambda\to-\infty$.

Es sei nun g ein beliebiges Element aus $\mathfrak{H}$; dann ist

$$|E_\lambda g|=|E_\lambda(g-f)+E_\lambda f|\leq|E_\lambda(g-f)|+|E_\lambda f|\leq|g-f|+|E_\lambda f|;$$

da nun die f, für welche Af Sinn hat, in $\mathfrak{H}$ überall dicht liegen, so kann man f und nach der Wahl von f auch λ_0, wie wir soeben bewiesen, so wählen, dass einerseits $|g-f|$, andererseits für $\lambda<\lambda_0$ auch $|E_\lambda f|$ und somit schliesslich $|E_\lambda g|$ beliebig klein ausfallen, woraus $E_\lambda g\to 0$ für $\lambda\to-\infty$ allgemein folgt.

Viertens gilt für $\lambda<\mu$, $\lambda\to\mu$ auch $E_\lambda\to E_\mu$. In Worten: E_λ ist als Funktion von λ nach links stetig.[17]) Wegen

$$(E_\mu f,f)-(E_\lambda f,f)=|(E_\mu-E_\lambda)f|^2$$

ist diese Aussage jener gleichwertig, dass die Form $(E_\lambda f,f)$ für jedes f eine nach links stetige Funktion von λ ist.

Um diese Behauptung zu beweisen, bemerken wir zunächst, dass zufolge der Monotonität von $(E_\lambda f,f)$ diese Grösse für $\lambda\to\mu$ einem Grenzwerte zustrebt und dass also

$$|E_\nu f-E_\lambda f|^2=(E_\nu f,f)-(E_\lambda f,f) \qquad (\lambda<\nu<\mu)$$

für $\lambda\to\mu$, $\nu\to\mu$ gegen Null geht. Somit geht $E_\lambda f$ für $\lambda<\mu$, $\lambda\to\mu$

---

17) Für die in der Folge zu entwickelnden Integraldarstellungen ist diese vierte Eigenschaft der Schar nicht von Belang. Sie könnte übrigens auch, wenn sie nicht von vorherein gesichert wäre, dadurch erreicht werden, dass man E_μ für alle μ durch $\lim E_\lambda$, $\lambda<\mu$, $\lambda\to\mu$ ersetzt.

gegen ein Element f' und die durch $E_{\mu-0}f = f'$ erklärte Transformation $E_{\mu-0}$ ist wegen $|E_\lambda f| \leqq |f|$ beschränkt und mit E_λ zugleich offenbar auch selbstadjungiert. Es gilt zu zeigen, dass $E_{\mu-0} = E_\mu$ ist; oder, indem wir die Bezeichnungen $\Delta = \Delta_{\lambda,\mu} = E_\mu - E_\lambda$, $\Delta \to \Delta_0$ für $\lambda < \mu$, $\lambda \to \mu$ einführen, so ist zu zeigen, dass $\Delta_0 = 0$ ist.

Sei f ein Element, für welches A Sinn hat. Dann hat A auch für $E_\lambda f$ und für $E_\mu f$, also für Δf Sinn. Durch Anwendung des Zerlegungssatzes auf Δf erhält man

$$(E_\mu(A - \mu E)\Delta f, \Delta f) \leqq 0; \quad ((E - E_\lambda)(A - \lambda E)\Delta f, \Delta f) \geqq 0,$$

woraus wegen $\Delta^2 = \Delta$, $E_\mu \Delta = (E - E_\lambda)\Delta = \Delta$ und der Vertauschbarkeit von A, E_λ, E_μ und Δ die Ungleichungen

$$((A - \mu E)\Delta f, f) \leqq 0; \quad ((A - \lambda E)\Delta f, f) \geqq 0$$

folgen. Aus diesen und aus $(\Delta f, f) \leqq |f|^2$ schliesst man, dass die Werte $((\mu E - A)\Delta f, f)$ und $((A - \lambda E)\Delta f, f)$ zwischen 0 und $(\mu - \lambda)|f|^2$ liegen, jedenfalls zunächst für alle f, für welche Af Sinn hat; da aber diese eine in $\mathfrak{H}$ überall dichte lineare Mannigfaltigkeit ausmachen, so folgt auf Grund einer im vierten Abschnitt gemachten Bemerkung, dass die Transformationen $(A - \mu E)\Delta$ und $(A - \lambda E)\Delta$ für ganz $\mathfrak{H}$ Sinn haben und beschränkt und ausserdem negativ resp. positiv definit sind, wie auch, dass für alle f aus $\mathfrak{H}$

$$|(A - \mu E)\Delta f| \leqq (\mu - \lambda)|f|.$$

Wegen der Vertauschbarkeit von A und Δ ist also auch speziell, wenn Af Sinn hat,

$$|\Delta(Af) - \mu\Delta f| \leqq (\mu - \lambda)|f|;$$

daraus folgt, für $\lambda \to \mu$, dass $\Delta_0(Af) = \mu\Delta_0 f$ ist. Da nun mit Δ zugleich auch Δ_0 mit A vertauschbar ist, so stimmt also die Transformation $\Delta_0 A = A\Delta_0$ für eine überall dichte lineare Mannigfaltigkeit mit der beschränkten Transformation $\mu\Delta_0$ überein und somit ist schliesslich identisch $A\Delta_0 = \mu\Delta_0$.

Es sei nun f wieder ein beliebiges Element aus $\mathfrak{H}$. Dann befriedigt also $g = \Delta_0 f$ die Gleichung $Ag - \mu g = 0$. Auf Grund der Aussage 3° des Zerlegungssatzes ist daher auch $E_\mu g = 0$. Andererseits folgt aus $\Delta E_\mu = \Delta$ und aus $\Delta \to \Delta_0$, dass auch $\Delta_0 E_\mu = \Delta_0$ ist und somit ist auch

$$\Delta_0 f = \Delta_0 E_\mu f = E_\mu \Delta_0 f = E_\mu g = 0,$$

d. i. $\Delta_0 = 0$, w. z. b. w.

Eine Schar E_λ von Einzeltransformationen mit den soeben bewiesenen inneren Eigenschaften, d. i. 1) Vertauschbarkeit; 2) $E_{\lambda_1} E_{\lambda_2} = E_{\lambda_1}$ für $\lambda_1 < \lambda_2$; 3) $E_\lambda \to 0$, E für $\lambda \to -\infty, \infty$; 4) $E_\lambda \to E_\mu$ für $\lambda < \mu$, $\lambda \to \mu$ nennt v Neumann eine *Zerlegung der Einheit,* Stone sagt dafür „*canonical resolution of the identity*". Die Beziehungen der Schar E_λ zur Transformation A sind durch ihre auf den Zerlegungssatz gestützte Definition gegeben; ich schlage vor, E_λ in ihrer Beziehung zu A in einer durch die Hilbertsche Terminologie nahegelegten Weise als die zur Transformation A gehörige *Spektralschar* zu bezeichnen.

Wir wollen nunmehr etwas näher auf die Beziehungen zwischen A und der zugehörigen Spektralschar eingehen. Wir haben schon im Laufe des soeben beendeten Beweises gezeigt, dass für $\varDelta = \varDelta_{\lambda,\mu} = E_\mu - E_\lambda$, wo $\lambda < \mu$, die selbstadjungierten Transformationen $(A - \mu E)\varDelta$ und $(A - \lambda E)\varDelta$ beschränkt und negativ resp. positiv definit sind; daher ist auch die selbstadjungierte Transformation $\varDelta A = A \varDelta$ beschränkt und es ist

$$\lambda(\varDelta f, f) \leqq (A\varDelta f, f) \leqq \mu(\varDelta f, f).$$

Für jedes ν zwischen λ und μ ist somit

$$|((A - \nu E)\varDelta f, f)| \leqq (\mu - \lambda)(\varDelta f, f) \leqq (\mu - \lambda)|f|^2,$$

also auch

$$|A\varDelta f - \nu \varDelta f| \leqq (\mu - \lambda)|f|;$$

setzt man noch $\varDelta f$ an Stelle von f, so erhält man

$$|A\varDelta f - \nu\varDelta f| \leqq (\mu - \lambda)|\varDelta f|. \tag{1}$$

Teilen wir nun die λ-Axe in abzählbar unendlich viele Intervalle (λ_k, μ_k) ein, die sich im Endlichen nicht häufen, und betrachten wir die entsprechenden Transformationen $\varDelta_k = E_{\mu_k} - E_{\lambda_k}$. *Die $\varDelta_k$ bilden ein vollständiges orthogonales System von Einzeltransformationen;* darunter verstehen wir, dass $\varDelta_k^2 = \varDelta_k$, $\varDelta_j \varDelta_k = 0$ für $j \neq k$ und $\Sigma \varDelta_k = E$, d. i. $\varDelta_1 + \varDelta_2 + \ldots \varDelta_n \to E$. Dies folgt unmittelbar aus den inneren Eigenschaften der Schar E_λ. Jedes Element f wird also durch unser System in abzählbar unendlich viele Komponenten $f_k = \varDelta_k f$ zerlegt; diese sind zueinander orthogonal, d. i. $(f_j, f_k) = 0$ für $j \neq k$ und es ist $\Sigma f_k = f$, d. i. $f_1 + f_2 + \ldots + f_n \to f$. Ferner ist

$$\sum |f_k|^2 = \sum (\varDelta_k f, \varDelta_k f) = \sum (\varDelta_k f, f) = (f, f) = |f|^2.$$

Nehmen wir nun an, dass Af Sinn hat und setzen wir $Af = g$; dann gilt analoges für g und dessen Komponenten $g_k = \varDelta_k g$.

Andererseits ist $g_k = \varDelta_k g = \varDelta_k A f = A \varDelta_k f = A f_k$, also ist

(2) $$Af = \sum A f_k.$$

Ferner ist die Reihe

(3) $$\sum |Af_k|^2 = |Af|^2$$

konvergent. Die Konvergenz dieser Reihe ist also eine notwendige Bedingung für die Existenz von Af; zeigen wir, dass sie auch hinreichend ist. Dazu bemerken wir zuerst, dass A für f_k immer Sinn hat, da ja die Transformation $A\varDelta_k$ beschränkt ist. Nehmen wir nun an, die Reihe (3) sei für ein gegebenes f konvergent. Da für $j \neq k$ $(A\varDelta_j)(A\varDelta_k) = A(\varDelta_j A \varDelta_k) = A(A\varDelta_j \varDelta_k) = 0$ ist, so sind auch die $Af_k = A\varDelta_k f$ zueinander orthogonal und daher

$$\left|\sum_{k=1}^{p} Af_k - \sum_{k=1}^{n} Af_k\right|^2 = \left|\sum_{k=n+1}^{p} Af_k\right|^2 = \sum_{k=n+1}^{p} |Af_k|^2 \to 0 \qquad (n<p;\ n\to\infty),$$

also konvergiert $Af_1 + Af_2 + \ldots + Af_n$ gegen ein Element f^*. Für jedes Element h, für welches A Sinn hat, gilt

$$(Ah, f_1 + \ldots + f_n) = (h, Af_1 + \ldots + Af_n);$$

für $n \to \infty$ geht die linke Seite gegen (Ah, f), die rechte gegen (h, f^*), so dass also $(Ah, f) = (h, f^*)$ und daher $Af = f^*$. Somit hat Af Sinn, w. z. b. w.

Jetzt wollen wir nunmehr von den Reihen (2) und (3) zu den entsprechenden Stieltjesschen Integralen übergehen. Nehmen wir zunächst an, wir hätten unsere Intervalleinteilung derart gewählt, dass die Länge der Intervalle eine gegebene positive Grösse δ nicht überschreite. Dann haben wir also nach (1) für alle k und für $\lambda_k \leqq \nu_k \leqq \mu_k$

$$|Af_k - \nu_k f_k| = |A\varDelta_k f - \nu_k \varDelta_k f| \leqq (\mu_k - \lambda_k)\,|f_k| \leqq \delta\,|f_k|;$$

daraus folgt die Konvergenz der Reihe

(4) $$\sum |Af_k - \nu_k f_k|^2 \leqq \delta^2 |f|^2$$

nebst der rechts angeschriebenen Abschätzung. Hieraus schliesst man erstens durch Heranziehung der Dreiecksungleichung

$$\left|\left[\sum_{k=1}^{n} |Af_k|^2\right]^{\frac{1}{2}} - \left[\sum_{k=1}^{n} |\nu_k|^2\,|f_k|^2\right]^{\frac{1}{2}}\right| \leqq \left[\sum_{k=1}^{n} |Af_k - \nu_k f_k|^2\right]^{\frac{1}{2}} \leqq \delta\,|f|;$$

diese Ungleichung zeigt, dass die Konvergenz einer der beiden Reihen

$$\sum |Af_k|^2, \quad \sum |\nu_k|^2\,|f_k|^2$$

jene der anderen nach sich zieht. Wir haben nun kurz vorher gezeigt, dass die Konvergenz der ersten Reihe eine notwendige und hinreichende Bedingung dafür ist, damit Af Sinn habe; dasselbe ist also auch für die zweite Reihe der Fall.

.. *Damit also die Transformation A für das Element f Sinn habe, ist es notwendig, dass bei jeder Einteilung der Zahlengeraden in Intervalle mit beschränkter Länge und für beliebige Wahl der Stellen ν_k auf den einzelnen (abgeschlossenen) Intervallen die entsprechend gebildete Reihe*

$$\sum |\nu_k|^2 |f_k|^2 = \sum |\nu_k|^2 |\varDelta_k f|^2 = \sum |\nu_k|^2 (\varDelta_k f, f) \tag{5}$$

konvergent ausfalle. Hinreichend dagegen ist schon die Konvergenz dieser Reihe in einem einzigen Falle.

Betrachten wir jetzt die Reihe

$$\sum (Af_k - \nu_k f_k). \tag{6}$$

Beachtet man, dass für $j \neq k$ neben $(f_j, f_k) = 0$ und $(Af_j, Af_k) = 0$ aus analogen Gründen auch die Gleichung $(Af_j, f_k) = 0$ besteht, dass also die Glieder der Reihe (6) zueinander orthogonal sind, so folgt aus (4) die Konvergenz der Reihe (6) und somit, sobald Af Sinn hat, mit der Konvergenz von $\Sigma Af_k = Af$ zugleich auch jene von $\Sigma \nu_k f_k$; ausserdem folgt noch die Abschätzung

$$|Af - \sum \nu_k f_k| \leqq \delta f. \tag{7}$$

Lassen wir jetzt unsere Einteilung derart variiren, dass die Maximallänge δ der Intervalle gegen 0 gehe, dann geht die Reihe (5) in das Stieltjessche Integral

$$\int_{-\infty}^{\infty} |\lambda|^2 \, d(E_\lambda f, f) \tag{8}$$

über; dabei bedingt die Existenz des Integrals (8) die Konvergenz der Reihe (5), während die Konvergenz von (5) — schon in einem einzigen Falle — die Existenz des Integrals (8) sichert. Damit ist also die Existenz des Integrals (8) eine notwendige und hinreichende Bedingung dafür, damit die Transformation A für das Element f Sinn habe.

Hat nun Af Sinn, so gilt für $\delta \to 0$ auf Grund der Ungleichung (7)

$$\sum \nu_k \varDelta_k f = \sum \nu_k f_k \to Af.$$

Mit einer naheliegenden Schreibweise kann man hiefür auch

$$Af = \int_{-\infty}^{\infty} \lambda\, dE_\lambda f \tag{9}$$

schreiben; übrigens folgt aus unserer Formel auch unmittelbar, dass

$$(Af, g) = \int_{-\infty}^{\infty} \lambda\, d(E_\lambda f, g)$$

ist.

Damit haben wir den folgenden Satz bewiesen:

Zu jeder beschränkten oder nicht beschränkten selbstadjungierten Transformation A gibt es eine Spektralschar E_λ (d. i. eine Schar von Einzeltransformation mit den oben aufgezählten inneren Eigenschaften) *derart, dass für alle f, für welche A Sinn hat, Af durch das Integral* (9) *dargestellt wird. Dabei hat A für jene und nur für jene f Sinn, für welche das Integral* (8) *konvergiert.*[18])

Der weitere Satz, nach welchem es zu jeder Spektralschar E_λ, die ohne vorherige Angabe von A durch die aufgezählten inneren Eigenschaften erklärt ist, eine selbstadjungierte Transformation gibt, zu der sie gehört (und die daher durch (9) erklärt wird),[19]) folgt durch eine leichte Umordnung desselben Gedankenganges.

Es sei schliesslich $F(\lambda)$ eine auf der unendlichen λ-Axe beschränkte und gleichmässig stetige, reell- oder komplexwertige Funktion. Dann konvergiert das dem Integral (9) analog definierte Integral

$$\int_{-\infty}^{\infty} F(\lambda)\, dE_\lambda f \tag{10}$$

für alle f aus $\mathfrak{H}$. Denn für zwei verschiedene Einteilungen in Intervalle (λ_k, μ_k) und (λ'_k, μ'_k) und die entsprechenden Transformationen $\Delta_k = E_{\mu_k} - E_{\lambda_k}$, $\Delta'_k = E_{\mu'_k} - E_{\lambda'_k}$ erhält man durch Heranziehung der gemeinsamen Unterteilung und der entsprechenden Einzeltransformationen $\Delta_{j,k} = \Delta_j \Delta'_k$, deren Produkt zu je zweien wieder gleich 0 und deren Summe wieder gleich E ist, und durch Anwendung einer für gewöhnliche Integrale geläufigen Schlussweise die Abschätzung

$$\left| \sum F(\nu_k) \Delta_k f - \sum F(\nu'_k) \Delta'_k f \right| \leqq 2\omega |f|,$$

[18]) Vgl. J. v. Neumann, a. a. O. [1]), Satz 36, ferner M. H. Stone, a. a. O. [2]), S. 424.

[19]) Vgl. [18]).

wo ω die maximale Oscillation von $F(\lambda)$ in den einzelnen Intervallen bedeutet; aus dieser Abschätzung folgt wegen $\omega \to 0$ definitionsgemäss die Existenz des Integrals (10).

Das Integral (10) definiert offenbar eine beschränkte und mit A vertauschbare Transformation, die wir mit $F(A)$ bezeichnen. Hat nun A für f Sinn und ersetzen wir in (10) f durch Af, so erhalten wir also

$$F(A)Af = AF(A)f = \int_{-\infty}^{\infty} F(\lambda) A\, dE_\lambda f. \tag{11}$$

Wir behaupten, dass unter der gemachten Annahme auch das Integral

$$\int_{-\infty}^{\infty} F(\lambda)\lambda\, dE_\lambda f \tag{12}$$

Sinn hat und ebenfalls gleich $F(A)Af$ ist. Betrachten wir nämlich die Differenz der beiden Integrale, d. i. das Integral

$$\int_{-\infty}^{\infty} F(\lambda)(A - \lambda E)\, dE_\lambda f \tag{13}$$

als Limes der Reihe

$$Sf = \sum F(\nu_k)(A - \nu_k E)\Delta_k f = \sum F(\nu_k)(Af_k - \nu_k f_k),$$

so folgt aus der Orthogonalität der Glieder der Reihe und aus der Ungleichung (4), dass

$$|Sf|^2 = \sum |F(\nu_k)|^2 |Af_k - \nu_k f_k|^2 \leqq (\max|F(\lambda)|)^2 \delta^2 |f|^2$$

ist und daraus folgt für $\delta \to 0$, dass das Integral (13) Sinn hat und gleich 0 ist. Somit hat mit (11) zugleich auch (12) Sinn und die beiden Integrale sind einander gleich, also ist

$$F(A)Af = AF(A)f = \int_{-\infty}^{\infty} F(\lambda)\lambda\, dE_\lambda f. \tag{14}$$

Es sei nun speziell $F(\lambda) = \frac{1}{\lambda - l}$, wo l eine nicht-reelle Zahl bedeutet. Setzen wir diese Funktion einerseits in die Integraldarstellung (10) von $F(A)f$, andererseits in die Formel (14) ein, so erhalten wir

$$F(A)f = \int_{-\infty}^{\infty} \frac{dE_\lambda f}{\lambda - l} \tag{15}$$

und

$$F(A)Af = AF(A)f = \int_{-\infty}^{\infty} \frac{\lambda}{\lambda - l}\, dE_\lambda f =$$

$$= \int_{-\infty}^{\infty} dE_\lambda f + l \int_{-\infty}^{\infty} \frac{dE_\lambda f}{\lambda - l} = f + lF(A)f,$$

also

$$F(A)(A - lE)f = f;\quad (A - lE)\,F(A)f = f.$$

Dies besagt, dass das Produkt der beiden miteinander vertauschbaren Transformationen $F(A)$ und $A - lE$, von denen die erste beschränkt ist, für alle f, für welche f Sinn hat und somit auch für ganz $\mathfrak{H}$ mit E übereinstimmt, dass also die durch (15) definierte Transformation $F(A)$ die Reziproke von $A - lE$ ist.

Damit haben wir eine Integraldarstellung der Reziproken von $A - lE$ für den Fall, dass l eine beliebige nicht-reelle Zahl bedeutet. Diese Darstellung behält ihre Gültigkeit auch für solche reelle Werte von l, in deren Umgebung E_λ konstant ist, d. i. für welche es Werte λ_0, μ_0 mit $\lambda_0 < l < \mu_0$ gibt derart, dass $E_{\lambda_0} = E_{\mu_0}$ ist. Man hat nur in diesem Falle unsere Integrale sinngemäss zu deuten, z. B. so, dass man die Funktion $\frac{1}{\lambda - l}$ in der Umgebung von l, d. i. in einem Intervall (λ_0, μ_0) von der angegebenen Art, durch 0 ersetzt; infolge des Verschwindens der den Teilintervallen von (λ_0, μ_0) entsprechenden Δ_k bleibt dann unsere Schlussfolgerung richtig, so dass also das Integral

$$\int_{-\infty}^{\infty} \frac{dE_\lambda}{\lambda - l} \tag{16}$$

die Reziproke der Transformation $A - lE$ darstellt.

Umgekehrt besitze $A - lE$ für ein reelles l eine beschränkte Reziproke C; wir zeigen, dass dann E_λ in der Umgebung von $\lambda = l$ konstant ist. Sei nämlich wieder $\lambda_0 < l < \mu_0$ und $\Delta_0 = E_{\mu_0} - E_{\lambda_0}$; ersetzen wir in (9) f durch das Element $C\Delta_0 f$ und bemerken wir noch, dass wir bei Berechnung des Integrals (9) λ_0 und μ_0 von Anfang an als Teilpunkte benutzen dürfen; daraus schliessen wir sofort, dass

$$\Delta_0 f = (A - lE)\, C \Delta_0 f = C \int_{\lambda_0}^{\mu_0} (\lambda - l)\, dE_\lambda f \text{ [20]}$$

und somit wegen $|\lambda - l| \leqq \mu_0 - \lambda_0$ auch

$$|\Delta_0 f| \leqq M_C (\mu_0 - \lambda_0) \left| \int_{\lambda_0}^{\mu_0} dE_\lambda f \right| = M_C (\mu_0 - \lambda_0)\, |\Delta_0 f|,$$

woraus für $\mu_0 - \lambda_0 < \frac{1}{M_C}$ das identische Verschwinden von $\Delta_0 f$, also $E_{\lambda_0} = E_{\mu_0}$ folgt.

Wir haben also den Satz:

Für alle nicht-reelle Werte von l, wie auch für alle jene reellen Werte, in deren Umgebung E_λ konstant ist, wird die Reziproke der Transformation $A - lE$ durch das Integral (16) *geliefert; für sonstige Werte von l besitzt die Transformation $A - lE$ keine beschränkte Reziproke.*[21])

Die weitere Erforschung der funktionalen Beziehungen zwischen *F(t)* und *F(A)* für integrierbare Funktionen allgemeiner Art, wie auch die analoge Untersuchung für den Fall mehrerer, auch abzählbar unendlich vieler, miteinander in gewissem Sinne vertauschbarer selbstadjungierter Transformationen[22]) lässt sich mit den hier benützten Hilfsmitteln, ja auch unter Vermeidung des Stieltjesschen Integrals, nämlich durch die schon in der Einleitung angedeutete sinngemässe Übertragung meiner alten Methode bewerkstelligen.

(*Eingegangen am 8. April 1930.*)

---

[20]) Ebenso folgt auch, dass das Ersetzen der unendlichen Integrationsgrenzen in den Integralen (8) — (16) durch endliche Grenzen λ, μ einer Anwendung der Transformation $\Delta = E_\mu - E_\lambda$ unter dem Integralzeichen resp. auf das durch das Integral definierte Element gleichkommt. Daraus geht wegen $\Delta \to E$ für $\lambda \to -\infty$, $\mu \to \infty$ sofort hervor, dass unsere in etwas ungewohnter Weise mittels unendlicher Reihen eingeführten Integrale auch durch Integration zwischen λ, μ und nachheriger Verschiebung dieser Integrationsgrenzen nach $-\infty$ resp. ∞ erhalten werden können.

[21]) Vgl. M. H. Stone, a. a. O. [2]), S. 423.

[22]) Vgl. J. v. Neumann, a. a. O. [14]); M. H. Stone, Linear transformations in Hilbert space III, *Proceedings of the National Academy of Sciences, U. S. America*, **16** (1930), S. 172—175.